Tests for non-additivity viewed as tests of the hypothesis of no interaction

A preliminary report

D. S. Robson

Abstract

Tukey's single degree of freedom test of the additive model

$$X_{ij} = \mu + \rho_i + \gamma_j + \epsilon_{ij} \quad , \ \{\epsilon_{ij}\} \text{NIID}(0,\sigma^2)$$

is viewed as a test against the alternative hypothesis

$$\mu_{ij} = u_i v_j \quad .$$

This view follows from the fact that Tukey's single degree of freedom completely accounts for the interaction sum of squares if and only if $X_{ij} = u_i v_j$. In an analogous manner, a single degree of freedom sum of squares may be formulated to completely account for interaction if and only if $\sqrt{X_{ij}} = u_i + v_j \geq 0$, thus providing a test against the alternative hypothesis of additivity on the square root scale. The procedure can be extended to other (specific) alternatives characterized by $X_{ij} = f(u_i, v_j)$.

Paper No. BU-334-M in the Bicmetrics Unit.

This research was supported by NSF - GB - 4502.

Tests for non-additivity viewed as tests of the hypothesis of no interaction

A preliminary report

D. S. Robson

BU-334-M Oct. 1970

Two factors A and B are said to be additive on the response scale X if there exist functions u(a), v(b) such that for every pair of levels (a,b) of the two factors,

$$E(X|a,b) = u(a) + v(b) .$$

If this additive model obtains on the X-scale then any non-linear change of scale will result in non-additivity and, conventionally, "interaction" would be said to exist on this new response scale. The existence of an additive response scale, however, implies that the two factors exert independent effects upon response, and non-additivity on any other response scale then represents an artifactitious type of interaction. Conventional usage of the term "interaction" in analysis of variance terminology is thus misleading in that the presence of "interaction" on a particular response scale does not imply interaction in the broader, lay sense of the word.

If we attempt to reduce the lay concept of interaction to a mathematical definition we are led to defining, instead, the antonym of interaction. Two factors are non-interactive if there exists a function f(u,v) such that

$$E(X|a,b) = f(u(a), v(b)) .$$

The functions f, u, v may depend on the response scale, but if such functions do exist on one scale then they exist on all scales. The existence of an additive scale is seen to be a sufficient but not a necessary condition for this type of "independence" (independent = non-interactive); thus even "non-additivity on all response scales" does not imply interaction.

A statistical procedure of testing for the presence of interaction as defined in this more general sense does not now exist, but the spirit is present in Tukey's "vacuum-cleaner" approach to partitioning interaction in the analysis of variance. In particular, his original one degree of freedom test for non-additivity may be construed as a test of the additive model

$$f(u,v) = u + v$$

against the multiplicative alternative hypothesis

$$f(u,v) = uv$$

since his single degree of freedom sum of squares completely accounts for the interaction sum of squares if (and only if) $X_{ij} = u_i v_j$, where X_{ij} is the observed response at the i'th level of A and j'th level of B (excluding those special cases where $\bar{u}$ and/or $\bar{v} = 0$). In such a case the "interaction sum of squares" of X_{ij} is clearly artifactitious as a measure of interaction in the general sense of the word. In application, a data analyst who discovers that Tukey's one degree of freedom accounts for virtually all of the interaction sum of squares will have made a significant discovery in his field of application by simultaneously demonstrating both the independence of the two factors and the validity of the multiplicative model. The importance of such a discovery would almost certainly outweigh that of all other information derived from the experiment.

In essence, Tukey's single degree of freedom procedure consists of testing the significance of the product-moment correlation between the observed residuals

$$e_{ij} = X_{ij} - \bar{X}_{i\cdot} - \bar{X}_{\cdot j} + \bar{X}_{\cdot\cdot}$$

in an r X c table and the f-predicted residuals

$$\hat{e}_{ij} = f\left(\hat{u}_i, \hat{v}_j\right) - \bar{f}\left(\hat{u}_i \cdot\right) - \bar{f}\left(\cdot \hat{v}_j\right) + \bar{f}\left(\cdot\cdot\right)$$

obtained by estimating $\left\{\hat{u}_i, \hat{v}_j\right\}$ from the equations

$$\bar{f}\left(\hat{u}_i \cdot\right) = \bar{X}_{i\cdot} \qquad\qquad \bar{f}\left(\cdot \hat{v}_j\right) = \bar{X}_{\cdot j} \ . \qquad (1)$$

Since the f-predicted residuals are constructed only from the row and column means, which are statistically independent of $\left\{e_{ij}\right\}$ under the null hypothesis of additivity with NIID $(0, \sigma_\epsilon^2)$ errors, then the single degree of freedom sum of squares

$$S^2 = \frac{\left[\sum\sum e_{ij}\hat{e}_{ij}\right]^2}{\sum\sum \hat{e}_{ij}^2}$$

is distributed as $\sigma_\epsilon^2 \chi^2_{1 d.f.}$ independently of the remainder

$$\sum\sum e_{ij}^2 - S^2 \sim \sigma_\epsilon^2 \chi^2_{[(r-1)(c-1)-1] d.f.} \ .$$

The construction of $f\left(\hat{u}_i, \hat{v}_j\right)$ from the row and column means is particularly simple in the case

$$f\left(\hat{u}_i, \hat{v}_j\right) = \hat{u}_i \hat{v}_j = \frac{\bar{X}_{i.}\bar{X}_{.j}}{\bar{X}_{..}}$$

but because f is non-linear the solution of equations (1) will usually entail iterative calculations. This is illustrated by the problem of testing additivity of X_{ij} against the alternative of additivity on the square root scale. Thus, if

$$\sqrt{X_{ij}} = u_i + v_j$$

$(u_i + v_j > 0)$ then, letting

$$\sigma_u^2 = \frac{1}{r} \sum_{i=1}^{r} (u_i - \bar{u})^2 \qquad\qquad \sigma_v^2 = \frac{1}{c} \sum_{j=1}^{c} (v_j - \bar{v})^2$$

we find

$$\bar{X}_{i.} = u_i^2 + 2u_i\bar{v} + \frac{1}{c}\sum_{1}^{c} v_j^2 \qquad \text{or} \qquad u_i = -\bar{v} + \sqrt{\bar{X}_{i.} - \sigma_v^2}$$

$$\bar{X}_{.j} = \frac{1}{r}\sum_{1}^{r} u_i^2 + 2\bar{u}v_j + v_j^2 \qquad \text{or} \qquad v_j = -\bar{u} + \sqrt{\bar{X}_{.j} - \sigma_u^2}$$

$$\bar{X}_{..} = \sigma_u^2 + (\bar{u} + \bar{v})^2 + \sigma_v^2 \qquad \text{or} \qquad \bar{u} + \bar{v} = \sqrt{\bar{X}_{..} - \sigma_u^2 - \sigma_v^2}$$

and

$$e_{ij} = 2(u_i - \bar{u})(v_j - \bar{v}) \quad .$$

Since

$$\frac{1}{r}\sum_{1}^{r}\sqrt{\bar{X}_{i\cdot} - \sigma_v^2} = \frac{1}{c}\sum_{1}^{c}\sqrt{\bar{X}_{\cdot j} - \sigma_u^2} = \sqrt{\bar{X}_{\cdot\cdot} - \sigma_u^2 - \sigma_v^2}$$

then

$$u_i - \bar{u} = \sqrt{\bar{X}_{i\cdot} - \sigma_v^2} - \sqrt{\bar{X}_{\cdot\cdot} - \sigma_u^2 - \sigma_v^2}$$

$$v_j - \bar{v} = \sqrt{\bar{X}_{\cdot j} - \sigma_u^2} - \sqrt{\bar{X}_{\cdot\cdot} - \sigma_u^2 - \sigma_v^2}$$

giving

$$\sigma_u^2 = \bar{X}_{\cdot\cdot} - \sigma_v^2 - \left[\frac{1}{r}\sum_{1}^{r}\sqrt{\bar{X}_{i\cdot} - \sigma_v^2}\right]^2$$

$$\sigma_v^2 = \bar{X}_{\cdot\cdot} - \sigma_u^2 - \left[\frac{1}{c}\sum_{1}^{c}\sqrt{\bar{X}_{\cdot j} - \sigma_v^2}\right]^2$$

and the latter two equations may be solved iteratively to obtain the former two. Letting $\hat{\sigma}_u^2(k)$ and $\hat{\sigma}_v^2(k)$ denote the solutions obtained in the k'th iteration then, starting with $\hat{\sigma}_v^2(0) = 0$, the two sequences

$$\hat{\sigma}_u^2(1) = \bar{X}_{\cdot\cdot} - \hat{\sigma}_v^2(k-1) - \left[\frac{1}{r}\sum_{1}^{r}\sqrt{\bar{X}_{i\cdot} - \hat{\sigma}_v^2(k-1)}\right]^2$$

$$\hat{\sigma}_v^2(k) = \bar{X}_{\cdot\cdot} - \hat{\sigma}_u^2(k) - \left[\frac{1}{c}\sum_{1}^{c}\sqrt{\bar{X}_{\cdot j} - \hat{\sigma}_u^2(k)}\right]^2$$

increase monotonically to their respective limits σ_u^2 and σ_v^2. As seen in the following numerical examples, convergence appears to be quite rapid.

Example 1: $X_{ij} = (u_i + v_j)^2$

X_{ij}

$u_i \backslash v_j$	0	2	5	mean
0	0	4	25	29/3
1	1	9	36	46/3
2	4	16	49	69/3
mean	$\frac{5}{3}$	$\frac{29}{3}$	$\frac{110}{3}$	16

e_{ij}

$u_i \backslash v_j$	0	2	5
0	$2\left(\frac{7}{3}\right)$	$2\left(\frac{1}{3}\right)$	$2\left(-\frac{8}{3}\right)$
1	$2(0)$	$2(0)$	$2(0)$
2	$2\left(-\frac{7}{3}\right)$	$2\left(-\frac{1}{3}\right)$	$2\left(\frac{8}{3}\right)$

$\sigma_u^2 = \frac{2}{3} = .6666 \cdots$

$\hat{\sigma}_u^2(1) = .47446$

$\hat{\sigma}_u^2(2) = .65933$

$\hat{\sigma}_u^2(3) = .66634$

$\sigma_v^2 = \frac{38}{9} = 4.2222 \cdots$

$\hat{\sigma}_v^2(0) = 0$

$\hat{\sigma}_v^2(1) = 4.10173$

$\hat{\sigma}_v^2(2) = 4.21734$

$\hat{\sigma}_v^2(3) = 4.22201$

Example 2: $X_{ij} = (u_i + v_j)^2$ fitted to $X_{ij} = u_i + v_j$

X_{ij}

$u_i \backslash v_j$	0	2	5	mean
0	0	2	5	7/3
1	1	3	6	10/3
2	2	4	7	13/3
mean	1	3	6	10/3

e_{ij}

$u_i \backslash v_j$	0	2	5
0	0	0	0
1	0	0	0
2	0	0	0

$$\hat{\sigma}^2_v(0) = 0$$

$$\hat{\sigma}^2_u(1) = .05128 \qquad \hat{\sigma}^2_v(1) = .35772$$

$$\hat{\sigma}^2_u(2) = .05783 \qquad \hat{\sigma}^2_v(2) = .35871$$

$$\hat{\sigma}^2_u(3) = .05785 \qquad \hat{\sigma}^2_v(3) = .35871$$

$$\hat{\sigma}^2_u(4) = .05785$$

$(\hat{u}_i - \hat{\bar{u}})$ \ $(\hat{v}_j - \hat{\bar{v}})$		$\hat{e}_{ij}$	
	$-.73721$	$.00741$	$.72980$
$-.30264$			
$.01685$			
$.28579$			

Note that the constraint $u_i + v_j \geq 0$, which is necessary if the square root transformation is to be appropriate, results in a unique solution to the estimation equations (1) applied to $f(u_i, v_j) = (u_i + v_j)^2$; otherwise there are 2^{r+c} solutions, and the procedure to follow if this constraint is removed remains an unresolved problem. A similar remark would apply, for example, to testing the model $f(u_i, v_j) = (u_i + v_j)^{-1}$ which has a unique iterative solution under the constraint $u_i + v_j > 0$.